



# Mirrors, Illusions and Epistemic Innocence

Roberto Casati

## ► To cite this version:

| Roberto Casati. Mirrors, Illusions and Epistemic Innocence. 2007. ijn\_00130684

**HAL Id: ijn\_00130684**

**[https://hal.science/ijn\\_00130684](https://hal.science/ijn_00130684)**

Preprint submitted on 13 Feb 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Mirrors, Illusions and Epistemic Innocence

**Roberto Casati**

*Institut Jean Nicod CNRS EHESS-ENS, Paris, and Università IUAV, Venice*

**Abstract :** Four major accounts of perception through a mirror are presented: two of them are unifier accounts (the entity one sees in the mirror is the same as the entity that gets reflected), and two are multiplier accounts (the entity seen is distinct from the entity that gets reflects: it is either a reflection, or a replica thereof). Most accounts have unwanted consequences, which manifest themselves by making perception through a mirror illusory one way or another. A unifier account is defended which reconciles metaphysical sobriety with epistemic un-innocence.

*Keywords: mirrors; perception; illusions*

Captain Hook stands before a mirror. What does he see? Hook has a hook in place of his right arm. But the person he sees in the mirror has a hook on his left arm. So it can't be the same person. Yet common sense says Hook sees himself; so it must be the same person. How can the dilemma be resolved?

The dilemma concerns what is seen "in" a mirror. What is the content of perception of a subject looking at an object reflected in a mirror, or to keep to a more neutral expression, a subject who uses a mirror as a visual aid? The difficulty in the answer lies in the fact that mirrors, like many epistemic artefacts, offer mediated access to reality, and such mediation appears to leave a trace (for example, in causing us to attribute different properties to whatever is perceived, as occurs with the change in position of the Captain's hook).

We will examine some accounts - not all will seem immediately plausible - that articulate the content of perception that occurs by means of a mirror. For there are various possibilities, grouped around the positions of unifiers (Captain Hook is what is seen) and multipliers (what is seen is something other than Captain Hook).

Ordinary parlance oscillates between various possibilities. At one extreme it indicates that one sees one's own *image* or own *reflection* in the mirror, while at the other it also allows that one looks at *oneself* in the mirror, in the sense that it is a real person that is seen, something in the real

world rather than an image. The implied ontology is different, and in the case of reflections the ontology needs to be clarified. Obviously it is not taken for granted that expressions faithfully represent intuitions, also because it is possible that here we are dealing with a semi-technical vocabulary, partially derived from optics (which uses ‘image’ in a mirror in a very specific sense of the term ‘image’). Nevertheless let us register a descriptive uncertainty to be investigated by considering some possible accounts that articulate an answer to the question of the nature of what one sees using a mirror.

An orthogonal distinction to that between seeing oneself/seeing an image is that of seeing oneself/seeing another *person*. More than by ordinary language, here the distinction is recorded by our behaviour, which highlights the possibility of an identification error.

We will also try to investigate what might be the cognitive reasons (the explanations of the intuitions) that make us choose one or another account of perception through mirrors.

## 1. The reflection theory

A first theory suggests that what is seen in the mirror is not the object that is reflected (Captain Hook) but a *reflection* of him. I don’t see Captain Hook but a reflection of Captain Hook, where it is presupposed that the reflection is a different item from that of which it is the reflection. This theory is in part faithful to common sense and in part distances itself from it. The commonsense element of the theory is the following. Let us say we see a reflection in the mirror in the sense that we say we see a shadow on the ground, separating thereby shadows as much as reflections from the objects of which they are shadow and reflection.<sup>1</sup> Shadows and reflections are not categorically homogeneous with the things of which they are shadows and reflections. A tidy way to express this intuition is to say one does *not* see the object, one sees *only* its reflection.

The notion of reflection however is very wide, and its primary accepted meaning includes the rays of light that reside on the surfaces of shiny objects and in which it is not always possible to recognise the object that caused them.<sup>2</sup> We limit ourselves here to the case where recognition is possible, so let us consider another account directly.

## 2. A variant: Images are seen in a mirror

People talk of seeing an *image* in the mirror; again in the sense where they mean to use the qualification to speak of a ‘mere image’. (There is also

---

<sup>1</sup> But cf. Sorensen 1999, for cases in which perception of a shadow is perception of an object.

<sup>2</sup> Cf. Cavanagh 2005. It is also necessary to ask if it is possible to specify what a reflection is without specifying what it is a reflection of.

a technical sense of ‘image’, a paramathematical meaning borrowed by optics, the sense of image as ‘mapping’, which will not be discussed here). To talk of images in a mirror, while allowing us to circumscribe those cases where *one recognises* something in the mirror, creates a further complication, insofar as the concept of image is associated with that of representational content. And in practice there is a sense in which one might think that mirrors were vehicles of representations: a sense in which it can be said that the mirror *represents* Captain Hook. It might be objected that mirrors are only “mechanisms” (loosely speaking) to reflect light, whereas the presence of representational content essentially must include reference to an intention. However, by introducing an appropriate intentional element one could answer the objection that sees in mirrors mere stages in the flow of information insofar as they do not involve intentions. I can have the intention to use a mirror to make you see what you look like, or to show you what is behind you. The representational content of the mirror in that case would be inherited from that of the intention governing its use. Therefore a causal-intentional account of image does not yet manage to exclude automatically that looking in a mirror is like looking at an image.<sup>3</sup> The relevant difference must be found elsewhere. The countercheck of the need to look elsewhere for the relevant difference between images and mirrors comes from the fact that some images too are obtained mechanically - such as photographs. The element of mechanical production is therefore not decisive.

One of the possible reasons for holding that actual images of things are what are seen in mirrors is a hypothetical analogy between the way things are seen in a mirror and the way images like pictures or photographs hanging on walls are seen. This analogy is supported by certain rational reconstructions of the origins of painting, and in particular of the representation of perspective, according to which one way for the artist to proceed would consist in tracing on the surface of a transparent screen (a

---

<sup>3</sup> This piece of evidence seems to motivate the arguments in U. Eco (1985) aimed at showing that mirrors are not semiotic phenomena.

window) the profile of the things visible on the other side of the window.<sup>4</sup> However we should beware of precisely this kind of rational reconstruction.

The unfeasibility<sup>5</sup> of the method just described is evident to anyone who has tried to carry it out. Parallax phenomena occur in windows that cannot occur in images. The smallest change in point of view causes a loss of alignment between the profile traced on the window and the profile of the thing to be represented. Indeed here lies the fundamental difference between images and windows. Windows do not function as images given that what is seen within a window changes according to adjustments of point of view, whereas what is seen by means of an image resists adjustments of point of view. But for this same reason mirrors do not function like images either, given that what is seen within a mirror changes in a way regulated by adjustments to point of view.

One can nevertheless extract from this discussion that if in need of a useful analogy to analyse the structure of the content of perception, mirrors function as windows rather than as images; and that in the case of mirrors it is still possible to fall back on the idea of image insofar as we are dealing with “windows” that show a reality that is different in some respects from that which exists beyond the surface of the mirror; by contrast with real windows, which in fact show the reality that exists on the other side of the window. Here we reconnect with the idea that mirrors present ‘mere’ images, and not reality in flesh and blood.

Up to now we have listed some aspects of mirrors that may cause intuitions to lean towards the assimilation of the content of perception when looking in a mirror, to content when looking at an image. Such assimilation

---

<sup>4</sup> Leonardo da Vinci, *Trattato della Pittura*, 402: “How the mirror is master of painters. When you want to see if your painting altogether conforms with the thing portrayed in nature, take a mirror, and make the live thing reflect in it, and compare the thing reflected with your painting, and consider well whether the subject of one and the other likeness conform to each other. Above all the mirror is to be taken as master, I mean the flat mirror, inasmuch as on its surface things have similarities with paintings in many parts; that is you see a painting done on a plane show things that seem in relief, and a mirror on a plane does the same; a painting is just a surface, a mirror is the very same; a painting is intangible, insofar as that which is round and distinct cannot be circled with hands, and mirrors do the same. Mirrors and paintings show the similarity of things surrounded by shadow and light, and one and the other appear well beyond their surface. And if you know that a mirror through features and shadows and lights makes things seem to stand out for you, and your having among your colours shadows and light more powerful than those of the mirror, of course, if you know how to compose them well together, your picture too will seem itself a natural thing, seen in a great mirror.”

If what I argue in this paper holds, the method Leonardo proposes is perfectly unusable: looking at something in a mirror does not introduce substantial differences compared to looking at it in reality. The passage from Leonardo’s treatise should be seen therefore more as a theoretical reflection on mirrors and images rather than as a series of practical pointers. And from a theoretical point of view it articulates - presenting them as obvious - some of the main points of a position that as we will see is not obvious.

<sup>5</sup> Casati 2000.

highlights the aspect of unreality that one would like to associate with what is seen in the mirror, but obscures certain important phenomenological aspects (in particular, as we will see in section 6, the important difference between the content of perception of mirrors/windows and images is given by the difference between the rules of co variation of content with movement of observer). We might say that this assimilation is based on ontological intuitions about the nature of what is seen, while its negation is based on phenomenological aspects.

### **3. Unifiers vs. multipliers: the simple unifier account**

An account that envisages that what is seen in the mirror is a reflection or an image of Captain Hook but not Hook himself is a *multiplier's* account: such an account multiplies the number of items relevant to the analysis of the structure of the content of perception through mirrors. As the supplementary entity is a reflection, call the theory a 'reflection multiplier' account. In opposition to multiplier accounts are deflationary or *unifier* accounts.

According to a first deflationary theory, it is not an image or a reflection of the object that is seen "in the mirror", but the *object itself* which is reflected, perhaps (but not necessarily) in the wrong place and with odd features. There are not two relevant items for the structure of perceptual content, but one.<sup>6</sup> Call this the 'simple unifier' account. On this account, the representational content of an episode of perception where a mirror is used as an intermediary is not substantially different from that of an episode where one is not used, aside from deformations incidentally introduced by and ascribable to the deviant causal chain.

In practice what we see in the mirror is a portion of the world: but, in virtue of the deviant causal chain, we represent that portion as if it were situated in a different place from where it is in fact located (for instance, on the other side of the wall in the hall where the mirror hangs, which we know to be the entrance to the neighbours' apartment not ours), and possessed of certain different properties from those it has (for instance, Captain Hook should have the hook on his right arm not his left).

At first pass, the content of perception of an object reflected in a mirror is therefore mistaken or illusory (in some sense that needs to be further specified) if the object has different properties from those that the content ascribes, as occurs normally. I see Captain Hook in the mirror. When I look in the mirror it seems to me that Hook is missing his left arm. But when I look at Captain Hook without using a mirror it appears his right arm is missing. The content of normal perception that has Captain Hook as its

---

<sup>6</sup> We have registered the expression "to see X *in* the mirror". A linguistic revision such as "to see X thanks to the mirror" captures an element of the unificatory account; as suggested above, it could be compatible with the expression "vedere allo specchio".

object and the content of perception that has as its object Captain Hook seen through a mirror therefore differ systematically in regard to certain descriptions, those featuring the terms ‘right’ and ‘left’.

How does the simple unifier account answer the dilemma outlined at the beginning of the paper? The account favours one of the horns of the dilemma as it subscribes to the intuition that what is seen in the mirror is the same thing that produces the reflection. But it must then explain how it is possible that things seen in the mirror have different properties from those that are being mirrored. The way out is to consider the content of perception through a mirror to be *illusory*. It is an illusion to have the impression of seeing the hook on the Captain’s left arm; it is an illusion to have the impression of seeing a space opening up before us when we look in the mirror.

#### **4. The replica multiplier account**

At the opposite end of the spectrum lies another multiplier account that aims to resolve the dilemma in a symmetrical way to that just set out: by denying (some aspects of) an attribution of illusion and at the same time denying the identity thesis. This multiplier account does not introduce items of a different type with respect to the objects that are mirrored, such as reflections or images, but introduces items of the same kind and postulates that they are (individual) objects distinct from those that are mirrored. The underlying intuition has been expressed in Borges’ story about Uqbar: mirrors are abominable because they increase the number of things.<sup>7</sup> So it is not Captain Hook I would see in the mirror, but a *replica* (which by definition is different from Captain Hook). This position also answers the mirror dilemma. It has the advantage of not postulating content that is mistaken about the properties of the object: the replica has exactly the properties ascribed by perceptual content determined by the perception of the mirror, and by definition that content is not mistaken about the properties it represents. The replica of Captain Hook is missing its left arm and I veridically see that it is missing its left arm. In fact, looking in the mirror would mean casting an eye over another possible world. The phenomenological comparison with windows returns here in another guise: it is as if a window had been opened in the wall the mirror is hanging on, revealing spaces and objects that are different from those actually on the other side of the wall.

But perceptual content has an *illusory* aspect as well. If I am not looking at Captain Hook but at a replica of him, how do I describe the behaviour - intentional or otherwise - of the replica? The replica winks. It does so at the same moment as Captain Hook (modulo the speed of light). Did the replica wink spontaneously? If we say yes, then we must consider

---

<sup>7</sup> “Then Bioy Casares recalled that one of the heresiarchs of Uqbar had declared that mirrors and copulation are abominable, because they increase the number of men.” J.L. Borges, 1940, “Tlön, Uqbar, Orbis Tertius”.

the impression that the *real* Captain Hook is responsible for the movement to be an *illusion*. Of course, when looking at ourselves in the mirror, we would find it particularly hard to regard as an illusion the fact that what happens “in the mirror” is completely determined by what happens in the world beyond the mirror. The replica account asks us to give up the basic intuition that says we *control* our reflection in the mirror.

## **5. The mirror illusion**

Both the simple unifier account and the replica multiplier account presuppose that there is an error in perception gained through a mirror; that mirror perception generates some form of illusion. To say that what is seen through the mirror is seen as incorrectly localised and is ascribed incorrect properties (deflationist theory) presupposes a mistake just as much as does attempting to save the veracity of content by postulating the correct perception of a replica with systematically different properties from the object of which it is a replica. One could nevertheless object that the error presents itself only if we accept that perception through a mirror is epistemically innocent, that is, produced without any awareness of the fact that we are dealing with mirrors. If such knowledge of the structure of the epistemic channel is integrated with the content, that is, if we consider epistemic innocence a myth, then some of the descriptive elements listed above are overturned, and the phenomenology of looking in a mirror must therefore be reinterpreted. Call this the un-innocent unifier account. Looking in a mirror is not (normally) like looking through glass, rather it is obtaining information about part of the world that sends to the mirror the light that the observer uses to perceive the scene. Captain Hook is not seen erroneously on the other side of the wall on which the mirror hangs, but on the same side of the room where the observer is located; and he is seen where he is precisely because it is known that a mirror is in use. Equally, he is not seen in the mirror with a hook on his left arm but with a hook on his right, that is, where in fact he has a hook. Note that a not indifferent advantage of this position is that it allows the instantaneous resolution of discussion about mirrors that ‘invert right and left but not top and bottom’ (Block 1974): the idea that an “inversion” occurred was an artefact of description: only abstracting from the fact that the reflective properties of mirrors are known (feigning epistemic innocence) could one maintain one had the impression that the hook had moved from the right arm to the left, and that therefore mirrors invert right and left. It is not content that is illusory; the illusion is the impression of having illusory content.

## **6. Homework for the un-innocent unifier account: a sensorimotor theory of reflections**

The un-innocent unifier account has on its side the generic advantages of ontologically parsimonious theories. Two things still need to be shown, however: in the first place, that multiplier accounts do not make virtuous use of the entities they postulate – that is, that such entities (images, reflections



and replicas) do not turn out to be indispensable or at least very useful in an account of the content of perception through mirrors. Secondly, that epistemic innocence is a non-unrenounceable feature of the phenomenological description of perceptual content.

On the first point (uselessness of postulated entities) : We saw in section 2 how the notion of image that is appealed to when it is held that seeing “in” a mirror is like seeing “in” an image is a notion that is in fact used to stand for (at least) three elements: the paramathematical use of the term ‘image’ (like ‘mapping’), the extenuating clause “only” (as in “one sees only an image in the mirror”), and the fact that images are representations. The fundamental difference between seeing in an image and seeing in a mirror was given by phenomenology: the robustness of what is seen in the image was contrasted with the transience of what is seen using a mirror. If I move from right to left relative to the picture of Uncle Sam, his threatening finger continues to point at me; but if I move to the left or right of an immobile Uncle Sam that I see through a mirror, I can escape his indicating gesture.

Nevertheless, phenomenology offers a hold to classificatory uncertainty. If we ideally bend a flat mirror until it becomes convex, we alter the parameters of coordination between our movements and changes in the visual image. The change in visual image is (relatively) minor compared to that which occurs when we move while looking at a flat mirror. At the limit, in cases of extreme convexity, the alteration is minimal: the reflections of light sources tend to remain anchored to the edges of shiny objects. Furthermore, an important “shrinkage” of reflected things can be noted in a convex mirror (think of the back of a spoon) compared to how they would be seen in a flat mirror.

This fact shows a phenomenological basis for what is seen in a mirror being attributed the character of image. We naturally tend to consider reflections in a convex mirror to be images insofar as they effectively behave like images (shrinkage and anchoring to the surface). The countercheck to this fact is suggested by a pin-hole camera<sup>8</sup> (which projects an image through a hole onto a translucent wall). The camera stands to a window as a convex mirror stands to a mirror: both the camera and the convex mirror tend to block the visual scene with respect to the movements of the observer.

On the second point (giving up epistemic innocence): Someone might want to hold that epistemic un-innocence, that is, the possibility of countenancing the structure of the epistemic channel when looking in a mirror, is not a feature of the phenomenology. Indeed, an objector would maintain, mirrors can fool us in a way that betrays our epistemic innocence.

---

<sup>8</sup> I owe the suggestion that a pin-hole camera is in a sense intermediate between image and window to Clotilde Calabi.

In particular in the case where they really seem to open up a space before us, a space in which we have the impression we can embark on an action. The reply to the objection is two-fold. In the first place, phenomena such as those described by the objector are marginal and surprising: it rarely happens that we are fooled by a mirror to the extent of taking the virtual space as a possible space for action - a space we would like to walk towards for example. Secondly, the marginality of these phenomena is tied to the fact that we learn to use mirrors successfully, to the point that epistemically un-innocent use of mirrors becomes second nature. In fact, we learn to ignore the space on the other side of the mirror; we learn to use rear-view mirrors (in right-driving countries, looking in the mirror, the car behind us does not seem to have the steering wheel on the right). This is also why attempts to 'stretch space' by putting big mirrors in narrow houses are made in vain. After a short while one becomes used to ignoring the 'virtual spaces' created by the mirrors. And for this reason we are surprised by simple tricks like illuminating something on the other side of a mirror by pointing a torch beam: for we do not normally consider the virtual space as accessible to action.

The un-innocent unifier account that embraces epistemic innocence at the phenomenological level is therefore the best candidate to solve the dilemma of Captain Hook. It solves it by denying one of the horns of the dilemma (we do not see the hook on the left arm of the Captain, but on the right arm, that is, where the Captain actually has a hook, by contrast with what is predicted by the simple unifier account), and accepting the other: what is seen in the mirror really is Captain Hook, as opposed to what the multiplier account predicts; we see neither a replica of him, nor his reflection, as against the predictions of the reflection and replica multiplier accounts respectively.

## 7. Morals

Some morals can be drawn from this discussion.

A. A first moral invites caution in the use of the terminology adopted to discuss 'virtual reality'. After all, mirrors provide an example of low-tech virtual reality.<sup>9</sup> In the case of mirrors, the 'virtuality' of what is seen through a mirror is nothing more than a turn of phrase to underline the fact that we are in the presence of a non-canonical way of collecting information about the world; but what is seen is the world itself, and the individuals who inhabit it.

B. A second moral concerns sensory-motor accounts of perception, which hypothesise that the content of perception is a function of knowledge of the type of movement which the perceiver can carry out in certain

---

<sup>9</sup> Cf. Ramachandran and Blakeslee 1998 for terminology. Note that Ramachandran's experiments use mirrors (which in this case offer a partial view) presupposing epistemic innocence.

contexts; even if there is now a consensus around a certain scepticism about their claim to be candidates for a general theory of perception (Block 2005), the intuition that guides them can be saved, at least locally. Sensory-motor principles work at the level of content to distinguish some contents of perception from others; in particular, here they have been useful to distinguish perception through a flat mirror and a convex mirror, and to show how the latter can be a step towards our notion of perception through an image.

C. A third moral concerns the use of notions such as image, which have a double life, both as ingredients of common sense and as elements of one or another account used to describe the world. Not to distinguish between the various meanings is to risk creating hybrid accounts that do not allow the correct classification of the phenomena to be described; and a descriptively inadequate theory has few chances of being explanatorily adequate.<sup>10</sup>

## 9. Bibliography

Block, N. 1974, “Why Do Mirrors Reverse Right/Left and Not Up/Down?”. *Journal of Philosophy* 71, 259-77.

Block, N., 2005, “[Review of Alva Noë, Action in Perception](#)”, *The Journal of Philosophy*, CII, 5, May 2005, 259-272.

Casati, R., 2000, *La scoperta dell'ombra*. Milano: Mondadori. Engl. Transl. *Shadows*, New York: Vintage.

Cavanagh, P. 2005, “The artist as neuroscientist”. *Nature* 434, 301-307

Eco, U., 1985 *Sugli Specchi*. Milano: Bompiani

O'Regan, J, Noë, A., 2001, “A sensorimotor account of vision and visual consciousness”, *Behavioral and Brain Sciences*, 24(5), 939-1011.

Ramachandran, V.S., Blakeslee, S., 1998, *Phantoms in the brain*. New York: William Morrow.

Sorensen, R. 1999, “Seeing Intersecting Eclipses”. *The Journal of Philosophy*, Vol. 96, No. 1, 25-49.

---

<sup>10</sup> Thanks are due to the attendees at the Gargnano meeting for insightful discussion, and to Milena Nuti for taking care of the editing.